NASA Technical Memorandum 89459

(NASA-TM-89459) WORKSHOP ON WORKLOAD AND TRAINING, AND EXAMINATION OF THEIR INTERACTIONS: EXECUTIVE SUMMARY (NASA) 40 p Avail: NTIS HC A03/MF A01 CSCL 05H

N87-25760

Unclas
G3/53 0083887

Workshop on Workload and Training, an Examination of Their Interactions: Executive Summary

Emanuel Donchin, Sandra G. Hart, and
Earl J. Hartzell

July 1987

NASA
National Aeronautics and
Space Administration

NASA Technical Memorandum 89459

Workshop on Workload and Training, an Examination of Their Interactions: Executive Summary

Emanuel Donchin, University of Illinois, Urbana-Champaign, Illinois
Sandra G. Hart,
Earl J. Hartzell, Ames Research Center, Moffett Field, California

July 1987

NASA
National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

EXECUTIVE SUMMARY - CARMEL VIII
Workshop on Workload and Training: An Examination of Their Interactions

Emanuel Donchin
University of Illinois
Urbana-Champaign, IL

Sandra G. Hart   Earl J. Hartzell
NASA Ames Research Center
Moffett Field, CA

INTRODUCTION

This report provides an executive summary of the Eighth Annual Carmel Workshop held in Carmel, California, from January 5 through January 10, 1986. The workshop was jointly sponsored by the Aerospace Human Factors Research Division and the Army Aeroflightdynamics Directorate, both at NASA Ames Research Center. The meeting was organized and chaired by Dr. Emanuel Donchin, Head of the Department of Psychology, University of Illinois at Urbana-Champaign.

The goal of this series of workshops has been to bring together a small group of researchers who represent two distinct disciplines that might benefit from sharing their findings and theories. The theme of the first seven meetings was the interaction between Cognitive Psychophysiology and other disciplines. However, in 1986, the series changed focus and brought together groups of experts in human performance in complex systems. The 1986 Workshop was predicated on the assumption that the fields of training and workload measurement have made considerable progress in the last two decades, but that each area of research has progressed in isolation from the other. It is clear that the level of training on any task may affect the workload associated with that task. It is also quite obvious that the workload encountered at different levels of training can have a considerable effect on the quality of training. We assumed that by bringing together leaders in the fields of workload and training, and providing them with the focus of considering the individual and joint relationships of their disciplines to challenging tasks (e.g., those imposed by advanced helicopters and the space station), this could have a substantial effect on integrating these disparate areas of research and enhancing their practical relevance.

It was clear from the quality of the discussions and the comments of the participants that the workshop succeeded in its goal. In addition, it clearly served as a forum where government officials, industry researchers, academicians, and government laboratory scientists could discover what each had to offer and what each needed. Considerable progress was made in examining deficiencies in the flow of information among the groups represented and in understanding their different perspectives, approaches, and goals. However, much was left to be discussed. Thus, a second workshop was scheduled for January 1987. Although the format will be the same, the topic will be broadened to include a consideration of the role of individual differences in workload and training. Although many of those who participated in 1986 will return, additional experts in the field of individual differences and selection were invited.

Format of the Carmel VIII Meeting

There were 36 participants from academia, government, and industry (Appendix A) in the Eighth Annual Carmel Workshop on Workload and Training. About one-third of the participants were experts in training, with an emphasis on those who had examined means of training operators to cope with complex systems. A second group of participants were experts in the definition, measurement, management, and analysis of workload. The third group of participants represented the Department of Defense and industrial organizations who are responsible for specifying, building, and managing advanced, complex systems. They must depend on the research foundation provided by the other groups and adapt it for their specific applications. In developing the list of participants, we sought diversity within each subdiscipline so that different classes of research and levels of application would be represented. Attendance was by invitation, and the participants were chosen not only for their expertise, but also according to their ability to contribute in this type of meeting.

The Workshop's program had three distinct stages. Excerpts from the letter of invitation (outlining the goals of the meeting) and the agenda are presented in Appendices B and C. During the first 2 days, 10 tutorials were presented that were intended to acquaint the participants with the activities within the domain of each group. These tutorials helped to define for the entire group the problems and considerations that each of its constituents considers critical and to specify the data base that each draws upon.

Following these tutorials, the participants were divided into four panels, each consisting of members from the three groups of participants. The panel received very detailed charges (see Appendix D) that consisted of questions related to a specific segment of the Workshop's main theme. The panels were given 2 days to develop a response to the charges and spent the time in intensive discussions. The participants reported that these working groups were very valuable in clarifying the nature of the interface between the groups represented at the workshop. The third and last phase of the workshop began on the fourth day. Each panel was allotted 3 to 4 hours to report their response to the charge to the entire group. In general, the time was split between presentations by members of the panels and active general discussions of the issues raised.

The meeting was adjourned following the panel reports. The tutorials and the reports of the panels were recorded in their entirety. The material is being transcribed and it is our intent to publish the proceedings of the conference within the next 18 months.

It can be said with considerable confidence that the Carmel VIII Workshop was an overwhelming success. Remarks received from some of the participants are included in Appendix E. It appears that a forum was created in which government, industry, and academic experts were able to examine each other's concepts and values in an unprecedented manner. This forum enabled a detailed examination of the theoretical foundations of training and workload, as well as the practical implications of these conceptual developments and the necessity for a more efficient and effective flow of information among the disciplines represented.

Rationale for the Workshop

Over the last decade, the concept of workload has been the focus of considerable attention. Hundreds of articles have reported empirical and theoretical treatments and many conferences have dealt primarily or solely with the concept, its measurement and prediction, and its applications. The concept of workload is intriguing from a theoretical perspective because of its intersection with research in the field of attention. However, much of the impetus for workload research has come from the practical need to evaluate workload during the design and operation of complex human-machine systems. For example, disputes about the number of crewmembers who will be needed to fly passenger jets and military helicopters in the future is often couched in terms related to the acceptability of the workload that will be imposed on the remaining crewmember(s), given a reduction in crew complement over current levels. Formal standards or even rules of thumb have not yet been developed to allow a user or designer of a new system to make this judgment without an empirical evaluation in a prototype of the target system.

The definitions of workload, as well as the measures used to quantify it, vary widely. Workload can be viewed as the demands imposed upon the limited information-processing capabilities of the human operator. Alternatively, other approaches emphasize the subjective experience of cognitive effort, the operator's level of arousal, and the physical, mental and emotional cost to the human operator of achieving the required level of performance. Gopher and Donchin considered workload to be a hypothetical, multidimensional, construct whose measurement must include behavioral, physiological, and subjective aspects. Indeed, workload measures in current use include overt response techniques (both primary and secondary task measures), subjective measures (in which the performer evaluates his or her workload), and psychophysiological techniques (in which various bodily responses, such as pupil size and brainwaves, are monitored).

Another area of research that is also experiencing a contemporary renaissance is training or skill acquisition. The performance of an advanced human-machine system is highly dependent on the expertise of its operators. Although much effort is put into simplifying systems and making them easier to operate, the enormous complexity of technologically advanced systems has created a substantial demand for well-trained operators. The period of training grows ever longer, and, in fact, never ends; continuing training is needed to maintain an operator's expertise. This increase in the length of training, the cost of the target systems, and the cost of training devices built to simulate them has created a situation in which small improvements in training will translate into savings of millions of dollars.

Another factor which contributed to the resurgence of research on training is the microcomputer revolution. Mainframe computers have been used for years to train a number of skills in many environments. However, the introduction of low-cost microprocessors into the world of simulation will allow a wide use of training manipulations that were previously impossible or prohibitively expensive. Presently, there are a few examples of inexpensive and powerful training devices (such as the tank gunnery trainer developed by Perceptronics). However, the new technology presents investigators with both new tools and new challenges.

The degree of operator training should, in theory, affect the workload experienced during task performance. It is also evident that the workload imposed by a task will affect the training process. Yet, research in these two areas has been conducted with minimal cross-contact. The separation between the two fields is unfortunate, because both research and application share similar difficulties. An example may be found in the requirement for task analysis common to both disciplines. Both face the difficulty of comparing results across studies in which different tasks were performed and different measures obtained. Both need objective performance criteria to evaluate the success of their procedures. And, both face the difficulty of extrapolating the results of previous studies to new tasks or environments.

The interaction between workload and training is rarely a topic of research and discussion. This silence is loud because it is clear that a connection between the two exists. Workload is clearly reduced when a task becomes highly practiced. For example, Schneider's description of his work on automatic processing illustrates well how the workload that an operator experiences is reduced with practice. On the other hand, the workload literature treats the influence of training as an external factor. Most studies focus on the workload levels achieved with asymptotic levels of performance and treat performance improvements as evidence for lower workload.

The influence of workload on training is equally, if not more, neglected. General practice holds that for instruction to be successful, workload should be maintained at an optimal level. Here again the importance of the topic is acknowledged, but the topic is then put aside with no further discussion. An additional source of confusion is the fact that in the past, when workload was introduced indirectly (as it was in research on massed versus distributed practice), a discussion of the influence of workload was limited to its impact on performance, not on learning. In fact, workload measures are rarely even obtained during the training stages of experiments designed to address workload.

The artificial separation between these two closely related domains is counterproductive. Both fields have grown, and the time is ripe for a scrutiny of the relationship between them. Crossing the lines, and discussing the relationship between training and workload is expected to be beneficial. Any merging of two bodies of knowledge is likely to advance both fields; this is particularly the case when the two domains have so much in common.

Focal Points for the Workshop

Helicopter Nap-of-the-Earth Flight- Of the many tasks and environments which confront human operators of complex, advanced technology systems, there are two that represent extremes of tasks and environments: single-pilot nap-of-the-earth (NOE) military missions flown at night in all weather conditions, and long-duration, multicrew, space station missions. These two tasks, and their environments, coupled with theoretical and applied issues related to training and workload were the foci of the workshop's deliberations.

The helicopter pilot's task in NOE missions might be defined as a rapid and continuous concatenation of goal directed anticipatory activity, motivated by stark fear, while attempting to coordinate and interact with other pilots similarly engaged and motivated. Quite aside from the obvious hostility of the NOE environment, flying in the vicinity of 40 mph only a few feet above the ground, produces relative velocities as great as those experienced by a fixed-wing fighter aircraft flying at hundreds of feet above the terrain. The NOE helicopter pilot is, however, constrained to fly down a tube of safety. Above some altitude, defined by surrounding terrain and vegetation, the helicopter is subject to radar detection and potential destruction by an enemy. The very terrain and vegetation that provide protection from detection constrain lateral maneuvering space, and unplanned intersection with the ground and flying too slowly close to the ground (which makes the helicopter vulnerable to hand-held arms or primitive missiles such as rocks and sticks) must be avoided. Further, the maximum speed potential of helicopters is limited; thus, helicopter pilots cannot climb rapidly to a high altitude or perform a rapid, high-g maneuver to escape a threat or give themselves time to think and replan the next segment of the mission.

In order to accomplish missions at night and in low visibility, pilots must wear light-intensifying goggles or use helmet-mounted, monocular displays of infrared imagery. While these visual enhancement systems are essential for any operations under these environmental conditions, their use places very high demands on the pilots. The field of view is restricted to about 40°, severely reducing essential peripheral visual cues. The images are displayed monochromatically, have limited resolution, and provide a limited forward field of view. In addition, helmet-mounted displays are presented monocularly, raising issues of binocular rivalry between the aided and unaided eye and the loss of stereoscopic cues. Thus, these systems impose high workload, even on trained pilots, and present a challenging training problem.

Generally, it is only when aircraft are in close proximity to the ground that new activities must be performed rapidly. Time to execute and plan actions during these periods is limited yet the consequences of errors can be catastrophic. Perhaps because of the time constraints, errors may even be more likely. These periods are brief and do not occur more often than twice per flight in most fixed-wing operations. With helicopters flying NOE, however, such periods are of long duration and occur frequently.

Another characteristic of NOE flight which is equally interesting and important is that navigation from one point to another is accomplished not by a straight line of flight, as in the case of fixed-wing aircraft, but by the circuitous routes dictated by the terrain, vegetation, and other obstacles. The environment and potential threats may be known only approximately in advance and it is difficult to remain spatially oriented. Therefore, one can add an element of opportunity and uncertainty to the mission; that is, the scenario of any NOE mission may change radically as a result of unexpected events or discovery of new opportunities. This imposes an additional requirement on the pilot: dynamic task reconfiguration in terms of goals and selection of systems to execute newly evolving tasks.

For the helicopter pilot, NOE flight presents only a relatively safe domain in which to complete a mission. It is characterized by a possibly intolerable number of competing demands for attention and action, and very little, if any, time which can safely be allocated to the planning and execution of command-control functions and corrective activities. Although this is a difficult state of affairs for the helicopter pilot, it represents a wonderful opportunity for researchers interested in understanding and solving problems of human behavior and performance in complex systems.

Space Station Missions- The Space Station is not a vehicle at all, but a facility operating in a near-Earth orbit. It is expected to be launched in the late 1990s; thus, much of what is envisioned is still conjecture. The facility will operate at altitudes somewhere between 250 and 300 n.mi. above the Earth while orbiting at a velocity of about 29,000 mph. In the first station, there will be a crew of eight who will stay in orbit for 90 days at a time. Later, as the facility is expanded, there may be as many as 30 people on board.

The Space Station will be built in orbit by assembling prefabricated elements delivered by a Shuttle fleet. The construction workers who assemble and build the Space Station will do so in hard suits rather than in the fabric suits used today, because of the hazards presented by space debris. A hard suit will be required not only for durability, but for mechanical advantage to construct the station. After the Space Station is completed, its crew will consist of scientists and engineers, for the most part, with additional crew-members charged with custodial, operational, and managerial responsibilities.

The Space Station will consist of five modules: one for the crew residence, one for biomedicine and animal studies (no human experimentation is planned), another for materials-processing research, and one each owned by Japan and a European consortium. The cost of building and launching the facility is estimated at $8 billion, with the United States providing $6 billion and the remaining $2 billion being split among the other nations--a true international undertaking. This orbiting facility will be the most expensive, well-thought-out and planned, high-technology facility ever built. It will be dominated by automated subsystems that will control and monitor the completely enclosed life support system on which the lives of the crew will depend.

The crew will likely work in two or three shifts, depending on the final size of the crew living quarters. For instance, two shifts would be elected if the crew quarters could support only four astronauts at a time with comfort; while half of the crew sleeps, the others would be on duty. This is called "hot bunking" and is a common practice in submarines. There are other similarities between the Space Station and submarines. They both represent long-term missions performed in confined spaces and hostile environments. The major difference is that submariners operate under military discipline and the crews are well-trained, hand-picked, and accustomed to working under a well-established chain of command. The Space Station, being an international civil operation, will not have the military model of command and discipline to depend upon. Furthermore, even though astronauts currently average 14 years of training, the future inhabitants of the Space Station can be expected to have no more than a few months of training. An interesting contrast between submarines and the Space Station is that submariners may not see the light of day for days at a time. Since the Space Station will make one revolution of the earth every 90 minutes, the crew of the station will see 16 sunrises a day or 1440 sunrises in a 90-day tour of duty. Thus, the problem of circadian rhythm must be addressed no matter what the work/sleep decision turns out to be.

The crew's survival will be totally dependent on the flawless operation of a completely enclosed and automated life support system. All air, water, and other fluids will be recycled. All odors that evolve, remain, and will be added to all previous odors. Further, there will be a 21-day turnaround time for each Shuttle flight up to the Station, and each round-trip flight is projected to cost about $120 million. If an illness or injury occurs, there will be a serious reluctance to make an unscheduled rescue flight. An emotionally disturbed crew member may present an even worse problem, as there is no provision for protective confinement.

From the crew's perspective, the environment represents complete isolation from the Earth on the one hand, yet complete absence of personal isolation from other crew members on the other. One cannot just go out and take a walk to get away from it all. For one thing, the cost of each extra-vehicular event is projected at $80,000 per hour. For another, the hard suit protective system is cumbersome and difficult to put on. Also, the noise from fans and other machinery is expected to be very high. Even now on the Shuttle, the noise level is 60 to 70 dB at all times. Finally, the problem of space sickness as a result of the microgravity environment is a real concern. In a 90-day mission it could occur at any time, not just in the beginning. If sleep disruption; long, fatiguing mission tasks; and a drab, unchanging, equipment-dominated environment are added to all conditions described up to this point, you have life on the Space Station. It is likely that at least mood and behavioral changes will be a common occurrence during a 90-day tour aboard the Space Station.

Thus, there are some remarkable differences between the two missions and environments, yet there are similarities as well. A Space Station crew member looking out a viewing port will perceive that the Station is moving very slowly relative to the Earth, even though it is orbiting at 29,000 mph. The helicopter pilot, on the other hand, flying at 40 mph a few feet off the ground will perceive movement that may feel faster than is comfortable. The primary similarities reflect the crew's dependence upon advanced technology and automation and the stresses induced by the missions and the environments. These two missions, and the vehicles that are expected to execute them, represent the extremes of what people will be asked to do in the future.

If these two systems and their environments represent the extremes of human involvement with complex systems, it is clear that there are problems for which we do not yet have all the needed information and solutions. The first step in the design of any advanced system is task and mission analysis. Estimates of training requirements, workload levels, and performance criteria begin with these analyses. Unfortunately, the actual tasks and missions and final designs are not yet available in any detail for either of these two vehicles. However, we do have some idea of the range and domain of the missions, environments, and tasks. Thus, when the complex interactions among training, workload, and performance issues are addressed in the context of these systems, it ensures some level of relevance and focus for the deliberations.

FUTURE WORKSHOPS

In the upcoming workshop (Carmel IX), one critical consideration will be added to the discussion of workload and training--that of individual differences. Far too much of our thinking is focused on the "average" individual, few of whom are ever encountered in practice. It is clear that the nature of training programs and their effects on, and interaction with, workload depend to a very great degree on the attributes of the individual operator. These issues were alluded to in the 1986 meeting, but there was insufficient time to dedicate the necessary attention to individual differences, and subject-matter experts were not present. This will be rectified in the 1987 meeting. Workload, training, and individual differences will be considered in the context of advanced-technology, single-pilot, helicopters. These vehicles present the most critical workload and training problems in the field of human factors today. Because workload levels will be so high and the margin for error so low, training must be very effective, workload must be evaluated very accurately, and the variability introduced by individual differences among operators may become critical.

APPENDIX A

EIGHTH ANNUAL CARMEL WORKSHOP

Workload and Training: An Examination of Their Interactions

January 1986

Final List of Participants

Michael G. H. Coles
c/o Dr. Marta Kutas
Department of Neurosciences M-008
School of Medicine
University of California, San Diego
La Jolla, CA 92093

Lloyd Neale Cosby
Perceptronics
1911 N. Ft. Myer Drive
Arlington, VA 22209

Mark Detweiler
University of Pittsburgh
Room 512, LRDC Bldg.
3939 O'Hara Street
Pittsburgh, PA 15260

Emanuel Donchin
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

Thomas Eggemeier
Engineering Psychology
Human Engineering Division
U.S. Air Force, AMRL
Wright Patterson AFB, OH 45433

John Ellis
NPRDC
Code 05
San Diego, CA 92152-6800

Keith Fender
Defense Chief of Staff for Personnel
DAPE-PSR
Washington, DC 20301

H. Thompson Fisher
Lockheed Missile and Space Co.
Org 53-33/Bldg. 580
1111 Lockheed Way
Sunnyvale, CA 94089

John D. Folley, Jr.
Applied Science Associates, Inc.
P.O. Box 1072
Butler, PA 16003

Richard Gabriel
Douglas Air Craft Company
Department E-71, Mail Station 35-36
3855 Lakewood Boulevard
Long Beach, CA 90846

Charles Gainer
Army Research Institute
P.O. Box 476
Ft. Rucker, AL 36362

Daniel Gopher
Faculty of Industrial Engineering
Israel Institute of Technology
The Technion
Haifa 3200, Israel

Peter Hancock
University of Southern California
Department of Safety Sciences
Los Angeles, CA 90089-0021

Sandra Hart
Man-Vehicle Systems Research Division
NASA Ames Research Center
MS 239-3
Moffett Field, CA 94035

James Hartzell
U.S. Army
Aero Flight Dynamics Directorate
MS 239-3
Moffett Field, CA 94035

Barry Kantowitz
Department of Psychology
Purdue University
West Lafayette, IN 47907

Art Kramer
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

Amir Mane
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

Neville Moray
Department of Industrial Engineering
University of Toronto
Toronto, Ontario
Canada M5S 1A4

Penny Sanderson
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

Walter Schneider
University of Pittsburgh
LRDC Building
3939 O'Hara Street
Pittsburgh, PA 15260

James Staszewski
Carnegie Mellon University
Department of Psychology
5000 Forbes Avenue
Pittsburgh, PA 15213

Michael Strub
Chief, ARI Field Unit
U.S. Army Research Institute
P.O. Box 6057
Ft. Bliss, TX 79916

Pamela Tsang
Man-Vehicle Systems Research Division
NASA Ames Research Center
MS 239-3
Moffett Field, CA 94035

Christopher Wickens
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

STAFF:

Barbara Hartman
Department of Psychology
University of Illinois
603 E. Daniel St.
Champaign, IL 61820

Earle Heffley
Department of Psychology
University of Illinois
603 East Daniel St.
Champaign, IL 61820

Wally Meyers
Department of Psychology
University of Illinois
603 E. Daniel St.
Champaign, IL 61820

OBSERVERS:

Vernol Battiste
NASA Ames Research Center
Mail Stop 239-3
Moffett Field, CA 94035

Anthony C. Beardsley
Grumman Aerospace Corporation
A05-25
Bethpage, NY 11714

C. Thomas Bennett, Major
U. S. Army
CMDR USA CDEC
Attn: ATEC-TP
Ft. Ord, CA 93941

Kent Hull
Man-Vehicle Systems Research Division
NASA Ames Research Center
Mail Stop 239-3
Moffett Field, CA 94035

Colin Kessel
Perceptronics
21111 Erwin Street
Woodland Hills, CA 91367

Robert Jay Shively
NASA Ames Research Center
Mail Stop 239-3
Moffett Field, CA 94035

Ben Tollison
Federal Aviation Administration
ASF - 300
800 Independence Ave., SW
Washington, DC 20591

Michael Vidulich
NASA Ames Research Center
Moffett Field, CA 94035

University of Illinois
at Urbana-Champaign

Department of Psychology
603 East Daniel Street
Champaign, IL 61820

College of Liberal
Arts and Sciences

APPENDIX B

November 8, 1985

Dear Colleague:

...It seems self-evident that the design of training systems need consider the workload associated with the tasks for which one trains. It is also reasonable to assume that the workload of a task changes as proficiency increases with training. Yet, one would hardly notice the relationship between the two areas of investigation from a perusal of the relevant literature. Studies of training barely note the variance introduced by varying workload, and students of workload tend to ignore the continually changing interaction between system and operator as practice on the task improves performance. The purpose of the Carmel Workshop is to examine the proposition that research on workload and training need be better integrated; to identify areas of interest common to investigators of workload and training, and to chart directions for research that will emphasize the interaction between the areas.

To provide a concrete framework for the discussions we shall examine the manner in which these issues arise in the context of the design of two specific systems, the Advanced Helicopter systems developed by the US Army and the Space Station. Both systems place extraordinary challenges on their designers and the design of training procedures is an important component of overall system design. The interaction of such training with the workload imposed by the systems will be of major concern for the workshop. Though we do, at the same time, hope to examine issues with due attention to the relevant theoretical issues...

...The original Carmel Conference series was sponsored by the Sloan Foundation, and its mission has been to bring together investigators from two areas of research where a common interest is likely but as yet undeveloped. The pattern we have found effective is reflected in the agenda for this meeting. The workshop's program is divided in three parts. We begin with a series of 10 tutorials. The tutorials are designed to acquaint members of each of the participating groups with the state of the art in the domain of the other groups.

Following the last tutorial the workshop breaks into 4 panels, each panel with 6-8 members, the membership of each is chosen to represent the various groups of participants. Each panel is given a fairly detailed charge. In the following day and a half the panels develop their responses to the charge. The charges to the panels for this meeting are enclosed with this letter as are the assignments of participants to panels. While meeting rooms are made available to the panels they are free to make any arrangements they find congenial and effective as long as they are ready with their response to the charge by the time it is needed.

In the last two days of the meeting the panels report to the workshop. We assemble again as one group and each panel is allotted half a day for its verbal presentation. It is left to each panel to design its presentation. The allocation of the panel's time is in the hands of its chairman. The most common arrangement has been to allot 20-30 minutes to each of the participants. However, some panels adopted different schemes. All are legal provided the charge is met and that ample time is left for discussion of the presentations...

...All participants are expected to attend the entire conference, beginning with the opening session and ending with the last panel discussion on Friday. Partial participation has a disrupting effect on the meeting and we have to be adamant that active participants, whose expenses are paid by the Workshop, attend the entire meeting. This rule does not apply to observers, who come at their own expense and who do not have a formal role in the Workshop...

...I am looking forward to seeing you in Carmel.

With regards,

Emanuel Donchin, Ph.D.
Professor and Head

ed/bjm

Enclosures: Agenda
Charge to Panels
Participants' List
Registration questionnaire

APPENDIX C

EIGHTH ANNUAL CARMEL WORKSHOP

Workload and Training: An Examination of Their Interactions

Agenda

Cognitive Psychophysiology Laboratory
Department of Psychology
University of Illinois at Urbana-Champaign

The La Playa Hotel, Carmel, California

January 5-10, 1986

Sunday, Jan 5, 1986

4:00-6:00pm Garden Room, Registration

8:00-9:00pm E. Donchin, University of Illinois

Structured Practice - A Path to Expertise. An examination of the role of part-training in the acquisition of expertise. The lecture will review the interim results of the Learning Strategies project.

Monday, Jan 6, 1986

9:00-10:00am James Hartzell, NASA Ames Research Center

The Training Challenge. A review of the challenges that are presented by the training needs of DOD. The importance of considering training, and workload, during procurement and design and the consequences of the failure to do so will be reviewed.

10:00-11:00am Tom Fisher, Lockheed Missile and Space Corporation
The Space Station

11:00-12:00am Charles Gainer, ARI Field Unit, Ft. Rucker
The Advanced Design Helicopter

These two tutorials, whose presenters are yet to be selected, will focus on two specific systems that currently present important workload and training challenges. In each case the system will be described with special attention to the skills in which the operators must be trained and the needs for workload assessment that characterize the design of these systems.

2:00-3:00pm Christopher Wickens, University of Illinois
The Concept of Workload

3:00-4:00pm Neville Moray, University of Toronto
The Measurement of Workload

4:00-5:00pm Sandy Hart, NASA Ames Research Center
Prediction of Workload

Three tutorials will present an overview of the state of the art in Workload Assessment. Wickens will present a theoretical analysis of the concept. Moray will be asked to review different aspects of Workload Assessment and Hart will consider methods for predicting, in advance, the workload that may be associated with a task.

5:00-6:00pm Meeting of Panel Chairmen, the Executive Suite

8:00-11:00pm Poster Session

This evening session has traditionally been set aside to allow participants to meet and discuss recent data from each other's laboratories. A room is set aside in the hotel, and participants bring in whatever data they would like to review on an informal basis with colleagues.

Tuesday, Jan 7, 1986

9:00-10:00am James Staszewski, Carnegie Mellon
Skilled Memory

10:00-11:00am Daniel Gopher, The Technion, Haifa
Task Analysis

11:00-12:00am Walter Schneider, University of Pittsburgh
The Mechanics of Training

These three tutorials will review current progress in the area of training and task analysis. The first tutorial will review the theoretical, and practical, status of the concept of skill, so that it can serve as a basis for discussions of skill acquisition. This will be followed by a review of the methods used for task analysis in different domains. The final tutorial will present a theoretical analysis of the training environment and the processes that are likely to play a role in training.

2:00pm Meeting breaks into four Panels. The Panels are scheduled to meet until the evening of Wednesday, January 8, 1986. Panel presentations are then made to the entire group according to the following schedule.

Wednesday, Jan 8, 1986

Panel Meetings all day

Thursday, Jan 9, 1986

9:00-12:00am Panel I: Measurement and Prediction Problems in Training and Workload

This panel will review the measurement of workload and the assessment of training outcomes. It will consider the degree to which the measurement of training should incorporate measures of workload. The measurements of workload include an assessment of proficiency and training achievement. The extent to which a synthesis between measurement procedures is possible and desirable will be examined.

2:00-5:00pm Panel II: The Interaction of Workload With Training

This panel will assess the evidence that the training process is, or is not, affected by the level of workload imposed by a task. The panel will review whatever is available in the literature on this topic. It will also examine the degree to which theoretical accounts of Skill Acquisition and of Workload predict the interaction between the two. Implications for further research will be examined.

This panel will also consider the effect that training has on workload. Of importance are such questions as the degree to which the workload associated with a task can be predicted from the level of mastery a trainee has reached. Indeed it would be important for this panel to examine the extent to which a training-free concept of workload is possible. An important aspect of the panel's work will be to examine the extent to which multiple resource models of workload imply a resource specific interaction between training and workload.

Friday, Jan 10, 1986

8:30-11:30am Panel III: The Effect of Training on Workload

Embedded training will receive special emphasis in this workshop. It is a particularly useful area in which to examine the interaction of workload and training, because it assumes that the trainee acquires skills during the actual performance of the task. Thus, workload may be at its peak while embedded training is to have its effect. The panel will examine the classes of embedded training currently in implementation, and those that come in planning. The panel will consider the need for and the measures whereby workload assessment need be included in the design of embedded training.

1:30-4:30pm Panel IV: Implications for System Design

This panel will work with the assumption that all other panels will have resolved the theoretical and practical issues associated with the interaction between workload and training. It will be this panel's task to place the theoretical discussions of the Workshop within the context of the designer of training systems and of man-machine systems. In particular, the panel will be asked to develop guidelines for a system design process that will take into consideration both workload and training considerations.

APPENDIX D

EIGHTH ANNUAL CARMEL WORKSHOP

Workload and Training: An Examination of Their Interactions

January 1986

Charge to the Panels

Emanuel Donchin
University of Illinois, Urbana-Champaign

I. Introduction

Following are tentative descriptions of the assignments for the four panels. As I said on the previous seven occasions, I am neither naive nor presumptuous enough to assume that this group of strong-willed individuals will, or should, follow my directives. The assignments are but general guidelines that identify a focus for the groups' discussions. I hope the groups will feel free to diverge from these guidelines wherever the discussion leads.

The times allotted for panel meetings are the afternoon of Tuesday, January 7 and Wednesday, January 8. Rooms will be set aside for the purpose. The panelists, however, are in control of the schedule and can meet at any other time. If the need arises let me, or Barbara Hartman, know and we shall arrange for rooms at other than the scheduled times. In the past some panels have benefitted from correspondence prior to the workshop. The time between this announcement and the actual meeting is short. However, the panel coordinators may wish to take action.

The panel coordinators, whose names are underlined in the panel membership lists will chair the panel meetings and the session of the workshop at which their panel's report will be presented. An important rule of these conferences has been that participants accept their panel assignment. While it is inevitable that some may find another panel more interesting, we have to abide by the present assignment. The arrangement of the panels is quite complex, as it has to satisfy many design criteria, and it would be very difficult to make any changes without substantial disruption.

The organization of the panels this time has proven somewhat difficult as there is considerable overlap between the topics of the various panels. In general, we see Panel I as considering the measurement problems associated with Workload Assessment and with the design of training systems. The following two panels examine in some detail the interaction between Workload and Training; one panel focusing on Workload and the other on Training. Finally, the fourth panel will examine the practical implications of this discussion with particular emphasis on System Design.

Panel I: **Measurement and Prediction Problems in Training and Workload**
Donchin, Gabriel, Kantowitz, Sanderson, Staszewski, Tsang

This panel is asked to review the procedures currently in use for the measurement of workload and for the assessment of training. This is a tall order as in both domains there is a vast literature, numerous techniques, and a host of controversies. Nevertheless, it is necessary for the subsequent discussions in the workshop that we examine, at the outset, the procedures that are being used for the quantification of Workload and of Training.

The panel is not expected to prepare an exhaustive review of all measurement procedures proposed, or tried, in these two domains. Rather, the panel is asked to develop guidelines for the evaluation of measurement techniques. Indeed, it will be a very useful outcome of the panel's deliberations if it can identify criteria for the evaluation of assessment techniques, criteria that are independent of the features of any specific measurement approach. The following questions may be addressed in a relatively general mode:

- To what extent is a measurement technique dependent on an explicit definition of Workload (or Training), and to what extent is the definition used consistently? In other words, what are the validity criteria for the measurement process? How consistent is the application of these validity criteria?

- If there is no explicit definition of validity criteria, what, if any, are the implicit criteria and how well are they applied?

- What, if anything, is the interaction between the definition of the concepts and the measurement procedures?

- What other assumptions are implied by the measurement procedure? What is the empirical support for these assumptions? When empirical support is lacking, how plausible are the assumptions?

- What are the consequences to the measurement process, its validity, and its reliability, of a violation of the assumptions?

- How do various attributes of the measurement process interact with subsequent use of the measures? Do the measurement attributes impose constraints on the statistical analysis of the data and on the interpretations that can be made of the outcome of the analyses?

- What can be said about the inherent reliability of the methods of measurement?

- What are the proper domains of application of the measures? It is possible to imagine substantial difference in the measurement goals when one is trying to characterize group differences, and the goals when one is in need of assessing the performance of an individual in specific circumstances?

These, and related questions, may be somewhat more contentious when applied to the measurement of Workload than to Training. The very definition of the concept of Workload is a matter for much debate. The panel will, no doubt, give attention to a comparative assessment of subjective ratings, overt performance measures and psychophysiological methods for assessing Workload. While we certainly do not expect this panel to resolve these thorny issues, the panel will serve the rest of workshop well if it does identify the key choices that need be made when adopting a method of measurement.

With respect to Training, matters are deceptively simple. If an operator is trained to perform a task, we can, it would seem, just measure the performance level as an index of training. This, after all, is the purpose of the "learning curve." Yet we are, in general, interested in predicting how well the trainee will perform over time, and in stressful circumstances. We would like to know how flexibly the material has been incorporated into the trainees' mental model. We need to know how often retraining will be needed and how many bad habits had been acquired in the course of training. Many of these very critical questions cannot be answered in a simple and direct manner by using observations on observed task performance. We hope the panel will attend to these issues as well.

Panel II: **The Effect of Workload on Training**
Bennett, Coles, Detweiler, Ellis, Gopher, Hartzell, Strub, Vidulich

This panel, and the one following, will confront the core issue of this conference. Both panels will consider the interaction between Workload and Training. However, we ask each of these panels to focus its attention on a different direction in the relationship. The present panel is asked to review the degree to which the design of training programs should consider the Workload inherent in the target task. We are distinguishing in this context between task "difficulty" and the workload associated with the task. A person who is entirely unversed in a task, a rank novice, does not necessarily operate at a high workload because the task, even when entirely unfamiliar, does not necessarily overload the person's resources. Other tasks do present a severe workload when unfamiliar, and as training proceeds the workload may either diminish or remain stable. This panel is asked to examine what is known about the effects of workload on training, as well as what is known about workload and training, and consider the degree to which knowledge regarding workload should affect the design of a training regime. Examples of questions we address to this panel are:

- Can the assessment of workload associated with a task be done separately from the assessment of the performance on the task?

- How should training change as a function of increases, or decreases, of the workload associated with a task?

- Assuming that practice is a mean for maintaining the ability to perform a task, how should practice be structured to take into account the differing level of task workload?

- Workload may be defined as the interaction between task demands, defined objectively, and the resources available to the operator for deploying in the service of the task. If this is the case, then individual differences in operator capacity determine the actual workload associated with the task. To what extent do we need to take such individual differences into account when designing training programs?

- Should workload measurement become a routine component of a well-designed training program? And, if so, how should such measurement be affected?

Panel III: **The Effect of Training on Workload**
Folley, Gainer, Hart, Hull, Kessel, Kramer, Schneider, Shively

This panel will examine the interaction between workload and training with a special emphasis on the effect that training may have on workload. In a sense it is the goal of research on both training and workload to bring about a reduction in the workload associated with performance as well as to reduce the amount of training that is required to achieve a given level of performance. As training proceeds the same level of performance can be achieved with a lower investment of the operator's resources and, by definition, the task's workload is diminished. The quality of a training regime can be assessed with reference to the degree to which it leads to a reduction in workload. It is to this aspect of training that we ask this panel to pay its attention. The class of questions addressed by this panel can be illustrated by the following questions:

- What is the relation between improved performance on a task and the reduction in its workload? Can performance improve in its quality without a related reduction in workload?

- Should workload reduction be recognized as a goal of training programs that is distinct from performance enhancement?

- If performance and workload are dissociated, are there ways to design training programs so that workload reduction will be achieved concurrently with the improved performance?

- Are there aspects of training that affect workload which are not directly related to level of performance? For example, one may conceive of the operator acquiring skills that enhance his or her ability to manage internal resources so that the cost in resources declines with training. How can such aspects be identified and provided for?

- As workload is a multifaceted phenomenon and as different methods of measurement focus on different aspects of workload, would the answer to the above questions vary with the aspect of workload that is being assessed in different circumstances?

One other class of issues this panel can address are those related to "embedded training." This label refers to training that the operator achieves while in actual performance of the task. Embedded training is a particularly useful area in which to examine the interaction of workload and training, because it assumes that the trainee acquires skills during the actual performance of the task. Thus, workload may be at its peak while embedded training is to have its effect. The panel will examine the classes of embedded training currently in implementation and those that are in planning. The panel will consider the need for and the measures whereby workload assessment need be included in the design of embedded training.

Panel IV: **Implications for System Design**
Battiste, Cosby, Eggemeier, Fisher, Heffley, Mane, Moray, Tollison, Wickens

The previous three panels are asked to examine the workload associated with tasks and the design training programs for the purpose of acquiring the ability to perform a task. The concept of "task" appeared as a primitive in the discussion of these three panels. However, tasks are in the main designed by people so that other people can work to satisfy the needs of yet another group of people. In other words, engineers design tasks so operators can satisfy management's goals. Once a task is designed and the operators are chosen, the system must assure that the operators are trained to perform the task and that the workload it presents is not beyond their capacities. What is all too often forgotten is that, in most cases, the time to minimize the workload associated with a task and the time to optimize the trainability of the task is at the time the task is on its designers drafting table.

This panel is asked to consider the manner in which the principles and the data discussed by the other three panels need be taken into consideration when actual tasks are being designed for use by real operators in functional systems. What, if anything, should a designer do to assure that workload is minimal? How should what we know about the way people can be trained affect the way we design systems? In what way can the design process incorporate the assessment of workload and the design of training programs?

An important issue that is embedded within this panel's domain is the complex of problems associated with the description and analysis of tasks. Can a designer identify, in an analytic manner, the skills that will be required for the performance of a task? Is it possible to predict workload analytically? What methods are available for task analysis and which method works best in which setting?

The panel is, of course, in somewhat of a quandary as it must begin its deliberations before the other panels have reported their conclusions. Yet the panel can no doubt predict the drift of the discussion. Furthermore, the panel, being the last to report, will have the advantage of having the last word. We hope that the panel can review for the workshop the interactions between workload and training as they appear from the perspective of the needs to design "real" systems.

APPENDIX E

The following pages are reproduced comments received from several participants of the Carmel VIII Workshop.

Lloyd Neale Cosby, Perceptronics, Arlington, VA

Keith Fender, Defense Chief of Staff for Personnel, Washington, DC

John D. Folley, Jr., Applied Science Associates, Inc., Butler, PA

Richard F. Gabriel, Douglas Air Craft Company, Long Beach, CA

Daniel Gopher, Israel Institute of Technology, The Technion, Haifa, Israel

Peter A. Hancock, Department of Safety Sciences, University of Southern California, Los Angeles, CA

E. James Hartzell, U.S. Army Aeroflightdynamics Directorate, Moffett Field, CA

Earle Heffley, Department of Psychology, University of Illinois, Urbana-Champaign

Barry H. Kantowitz, Department of Psychology, Purdue University, West Lafayette, IN

Neville Moray, Department of Industrial Engineering, University of Toronto, Ontario, Canada

Penny Sanderson, Department of Psychology, University of Illinois, Urbana-Champaign

Jim Staszewski, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA

Michael H. Strub, U. S. Army Research Institute, Ft. Bliss, TX

Pamela Tsang, NASA Ames Research Center, Moffett Field, CA

Michael Vidulich, NASA Ames Research Center, Moffett Field, CA

Remarks of Participants in Carmel VIII

Note: The remarks are essentially presented in their entirety. Minor editing was made to remove extraneous material and to eliminate personal responses. All references to individuals have been changed so that the person's full name appears in the text.

Lloyd Neale Cosby, Perceptronics, Arlington, VA

Thanks again for a most informative workshop. It was a highly productive week for me. From my viewpoint, the agenda cut to the heart of the root issues of military training.

My conclusions are the same as the final hour of the workshop; namely, we have a lot more to offer than we are prepared to admit on individual training and workload, but need to do a lot of work on the collective side. I tried to visualize this point by showing Chris Wickens' chart and Keith Fender's chart in my summary, i.e., the military requirement and academic status of where we are today.

Keith Fender, Defense Chief of Staff for Personnel, Washington, DC

The workshop was useful from a number of points of view, all of which are the basis for next year for further exploitation. The government was able to explain their position as the ultimate consumer, specifically addressing what decisions needed to be made and what tools are needed to support those decisions. The government also had an opportunity to view the wares of basic research projects to better judge their value.

Industry exercised their role of being between furnishing a product to the government and applying tools developed from basic research efforts. The research community benefitted from exposure to the problems of government and industry who are underwriting basic research's work (more or less). There was also an obvious benefit from presenting papers and the exchange of ideas that I will not discuss, as others are more qualified.

There is a need for government to articulate where research needs to be done and then ensure a stable funding base for that effort. It is clear that despite the lack of a universal theory of workload, useful tools are being created to measure cognitive workload. Further, there is exciting work being done in basic areas that can potentially spin off applied projects.

John D. Folley, Jr., Applied Science Associates, Inc., Butler, PA

1. Systems are still designed and built without adequate human factors engineering--despite the specifications and regulations requiring it.

2. Workload results from the way the system is designed--the inputs to the operator, the outputs required of the operator, and the implied cognitive processes between.

3. The point has probably been reached in some systems where training, no matter how extensive or expensive, cannot produce operators of sufficient skill; the task demands are beyond the capability of the available population or perhaps beyond the capability of anyone.

4. Some university research is definitely directly applicable to so-called "real-world" problems. I think of two in particular:

   a. Danny Gopher's work with the "space fortress" task.
   b. Walter Schneider's work on troubleshooting.

The feature that makes research work applicable to real-world tasks is that they are complex enough and a given experimental subject works at the task for a realistic period of time.

5. In contrast, some of the research has little applicability. I think of the report on "massed vs. distributed" practice on a pursuit rotor, with trials and intervals measured in seconds.

6. The conference was very worthwhile in my opinion. Its value could have been further enhanced, I believe, by a final plenary session at the end, the purpose of which would have been to reach a consensus on some conclusions and on some guidance for researchers, for those who use the results of the research, and for those who plan similar future conferences.

Richard F. Gabriel, Douglas Air Craft Company, Long Beach, CA

In response to your request concerning a brief personal summary and conclusions regarding the Workload and Training Workshop held in Carmel in January 1986, the following is offered.

Personal Impact. In general, I thought the week's workshop had a direct impact on me, specifically:

- Several useful discussions on the need-role of theory even in an applied activity.

- Bringing me at least partially up to date in some of the more recent training research.

- A useful review of current mental workload assessment research approaches.

- I gained new insights into the relationship between training and workload.

- Fruitful informal exchanges with authorities in these fields and personal contact that will make information exchanges more productive in the future.

Group Impact. Judging by the comments of others at the workshop with whom I conversed, I would offer the following general observations:

- The group, specifically those from academe and NASA, gained an improved understanding of the needs for measurement by government and industry.

- The entire group, through panel IV's efforts, gained insight into government and industry's problems in the acquisition process.

- There was useful organization provided by individual speakers of a number of relevant topics such as types of man-machine models, types of validity characteristic of each of the workload measures, etc.

- A better understanding of how government, academia, and industry might more productively pool their efforts.

In conclusion, I sincerely believe the workshop was a success. I have briefed several of my staff here at Douglas Aircraft on the activities, and have had a positive reaction from them regarding the format and content of the meeting.

<u>Daniel Gopher, Israel Institute of Technology, The Technion, Haifa, Israel</u>

Assessment of operator workload and the development of training schedules for complex tasks are two key problems of concern to contemporary Human Factors Engineering. While the two are basically separate, there are important areas of overlap between them that should be brought forward and evaluated. The Carmel meeting was one of the first of its kind in being devoted to the overlap between the topics. It was also somewhat unique in its assembly of participants that included a mixture of academic researchers and field professionals from industry and government operational settings. The combination of people and the selection of topic proved to be fruitful in generating different angles of view and novel ideas.

The general claim that proper training can act to improve performance in high workload tasks was unanimously accepted. Moreover, a large core of experimental evidence was cited by various participants to show that the gain in a few hours of training may be substantial. Two major directions of training have been identified. One type leads to a reduction of the overall load of a task, by automatizing as many as possible of its subelements. The second type influences performance by teaching subjects to cope more efficiently with the demands of high workload situations, in which the state of high load is permanent concommitant. Procedures to automatize task performance that draw upon a model of automatic and controlled processes were presented by Walter Schneider. They build upon the identification of the consistent elements that exist in every complex task and the training of subjects to perform them automatically without calling upon attention-demanding controlled operations.

Training schedules for efficient coping with the demands of tasks in which high workload is a steady-state property, was discussed by Daniel Gopher. Tasks in this category are usually complex tasks such as air-to-air dog-fight, or performance in a Command and Control (C3) environment, when an exact performance criterion does not exist and the general motivation is to excel. Under those conditions, uninstructed practice tend to lead operators to develop suboptimal coping strategies. The training approach proposed by Gopher is based upon a multiple resource model of attention. It directs individuals through priority changes of components, to learn the efficiency of alternative response strategies, thereby leading to the development of an individualized, task optimal arsenal of performance strategies to cope with task demands.

The discussion of training approaches was given a deeper meaning and a broader perspective by the preceding and following arguments on the role of theory in applied work, a review of current theoretical models of attention and resources, and an examination of the prevailing methods for task analysis. The general tone of academic and scholarly discussion was continuously perturbed by the harshness of real-life issues interjected by the industry and field people. The questions were generally referenced to the LHX and the Space Lab projects. Most revealing in this regard was the simulation game run by the system design discussion panel chaired by Amir Mane. The difficult, nonetheless important, role of human factors in such projects emerged with all its colors during this game. It became immediately evident how lacking our state of knowledge is in the area of workload and training, compared to the requirements and expectations of the field. However, it was also clear that human factors has a contribution to make, and that its chances to be substantial increase exponentially the earlier the involvement of the human factor specialist is in the project and the closer the project is to its starting point.

Peter A. Hancock, Department of Safety Sciences, University of Southern California, Los Angeles, CA

Preamble. The Carmel Conference on Workload and Training was founded upon the assumption of the possibility of a fruitful interchange of information between the described topic areas. An additional dimension of the Conference was the spectrum of participants who represented the concerns of government, industry, technical consultants and academe.

Benefits. The most obvious success of the Conference was in the affirmation of the validity of the assumption upon which it was predicated. Both the theoretical and practical necessity to link the two areas of investigation was made abundantly clear in the summary represented in the panel reports. Several avenues through which such integration might be achieved were addressed. The principal theoretic effort, predicated upon Schneider's automatic and controlled processing distinction, promised to offer fruitful possibilities. The potential utilization of alternate concepts (e.g., attentional resource utilization efficiency) were largely unexplored in formal meetings due to time restriction. Of the many strategic manipulations through which training may be instantiated, the part-whole approach received greatest airing. However, the absence of comprehensive data (with the exception of that presented for the space fortress game) prevented a thorough analysis of the potential for differentiating common performance "tasks" into behaviorally meaningful subunits. What functional impact that the fractionation of whole tasks exerts upon perceived workload and how subsequent task reassembly affects performance, learning, and workload became clear necessities for further research effort.

The problems associated with specific operational systems (e.g., LHX) were distilled in the comments of many participants. These reinforced the necessity for a comprehensive theoretical framework from which to extrapolate to individual circumstances as championed by Kantowitz. This may be considered as premature in that the theoretical underpinnings of both workload and training individually are as yet uncertain. However, a primary benefit of the Conference was in emphasizing potential points of integration and mutual insight, albeit through the use of some common mediational concept (i.e., in this case attention).

Future Directions. It is clear that this, hopefully inaugural, meeting pointed to the need for continuing interchange. Future efforts might consider a formalization of a common theoretic structure and also alternate strategies for manipulation of workload and training (e.g., the presence or absence and form of information on knowledge of performance (KP) and knowledge of results (KR) in training and workload interaction. An alternate theme might consider computer-based methods to aid in the optimization of training and workload coaction. The "mix" of participants is a feature of the meeting which should be perserved in that some of the most important insights came from the appreciation of problems as perceived through the eyes of others. This was most clearly exemplified in the presentation of Panel IV.

Summary. The meeting confirmed the necessity for interaction between interests in Workload and Training and emphasized the need for continued formal interchange of ideas and insights.

Earle Heffley, Department of Psychology, University of Illinois, Urbana-Champaign

My task was to summarize the recommendations of the panel on implications for system design. The panel generated 14 recommendations.

1. Workload (and training) should be considered early in the development cycle--that is, during the design phase.

2. We should work toward a standardized methodology for workload analysis, applicable to design phases.

3. Human factors should be given significant weight in system evaluation.

4. The academic community

   - Should not underestimate what it knows and can contribute,

   - Should not underestimate the value of relatively simple guidelines,

   - Should not oversell its existing collective wisdom as exact standards or precise numeric indices.

5. Good task analysis is absolutely essential; it should be

   - thorough with respect to components of human information processing,

   - comprehensive with respect to major missions/objectives/ conditions.

6. Cognitive components (system management, high-order pattern recognition) of system operation should be given increased emphasis.

7. Consideration of overload conditions (weather, faults, battle) should be included in all phases of the cycle--design, evaluation, and training.

8. Components performed on a collective basis should be considered in addition to components performed independently by individuals. We should not ignore analysis and training of strategy components, such as tasks performed by a group commander.

9. Training analysis, in early phases and throughout the cycle, should consider skill maintenance as well as initial learning.

10. Strong emphasis should be given to validation of methodologies and principles.

11. We should develop/enhance/validate computer programs (and other tools) for the analysis of workload.

    - Based upon models of human information processing, the knowledge of experts, etc.

    - Input: detailed timelines from task analysis

    - Output: workload profiles.

12. We should develop/enhance/validate computer programs (and other tools) for the analysis of training requirements.

    - Based upon analysis/synthesis of information from existing training databases (for example, commercial and military aviation).

    - Input: task analysis in terms of skill components.

    - Output: summary of initial training and skill maintenance factors (repetitions requires, duration of training phases, etc.).

13. Agencies should consider development of "generic" simulators that could be used to study/validate guidelines/principles/ methodologies.

14. We need a better flow of information between industry and universities: student internships, cooperative relationships, etc.

E. James Hartzell, U.S. Army Aeroflightdynamics Directorate, Moffett Field, CA

The following is in response to your request for commentary on the Carmel VIII workshop. These notes represent a heavily edited version of the introduction to a talk I would have given at the workshop had I not had laryngitis. Much of the remaining portions of the talk are included in the Executive Summary.

The Training Challenge

I was listed on the agenda to talk about the Training Challenge. This presented me with a challenge and a dilemma of my own. I was supposed to address the importance of considering training and workload during the design and procurement process within the DOD. My dilemma derived from the fact that the armed services by and large don't design, let alone build, anything; they write specifications and industry designs and builds whatever it is to these specifications. DOD procures the end item and then compares the product to the original specifications to determine its suitability. With increasing frequency DOD may even contract for the writing of the specifications. Thus, what DOD really does is to provide the need for things, expressed and manifested in the form of a new mission or mission requirements.

The remarkable advances in technology have not only captured the imagination of the technologists, but also that of the general population. However, complex and sophisticated systems are often developed without adequate reflection on the consequences to the human interaction with the system or its technology. It seems that technology can do anything man can think of. However, just because it is possible to develop something does not compel or justify one's doing so. Rampaging technology may be leaping over serious scientific issues and leaving behind theoretical holes that may one day soon bring us upon desperate times, financially and technically.

Whereas technology is evolving seemingly unencumbered, man has reached the end of the evolutionary trail. We are stuck with limits on reaction time, visual acuity, intellect, memory, and so on. The evolution of man and the evolution of technology are on divergent courses. To make matters worse, advanced technology manifests itself in the form of AUTOMATION.....a word that inculcates the attributes of advanced technology, which implies a promise to make life easier and more pleasant for humans. The idea conjures up the vision of little Expert Systems running around taking care of all manners of things. I am being overly critical of the idea to make a point. It is as though we believe that if a task is difficult it should be automated away. The assumption is that the workload goes away with it and, with less tasks to do, there should be less training required. Let the pilot become a manager, a monitor of automated systems. However, humans are very good at psychomotor tasks and not very good as monitors. Not only have we come to the end of our evolutionary trail, we may have come to the end of our piloting trail. Since it is unlikely that advances in technology will wait for us to catch up, what can scientists and researchers in the field of human function and behavior do to help?

I see the challenge to be at least two-fold: We must first elucidate what is known about human function and behavior while interacting with complex and advanced technology systems, to squeeze the last drops of proficiency and efficiency out of our remarkable, but limited and fixed, human capabilities. The next problem is to make the technologists and engineers aware of the consequences of their design decisions and applications of technology, as they impact the limitations and capabilities of the human interacting with the technology in terms understandable to them. Making them aware by fiat, edict, regulation standards, and thinly veiled threats is doomed to failure. Heightening their awareness of human issues in design through educational methods and media is a noble effort, and must be done. Unfortunately, this often serves only to give them access to the "buzz words" needed to convince the regulator that the "regulatee" has complied with the regulations. We must provide the designers, technologists, and industrialists with methods, measures, and information in terms understandable and usable to them or we have only placed another costly layer of regulations on the problem; a band-aid on a festering national sore when systemic medication is clearly indicated. Methods and measures are a difficult enough problem, but information is the most critical issue of all. Designers have a long history of going to a handbook for guidance. More recently, computerized data bases may solve the problem of conveying information to the designer. Some scientists have even attempted to express their data in engineering terms. The shortfall here is that isolated facts and data extracted from a data base deny the user of the information and insight into the relationships and interactions among the facts and data. As systems have become more complex, there has been a combinatorial explosion of significant interactions and inferences which are critical in the design process.

My final point, then, is that one of the new frontiers in our science should be the development of methods and means of understanding the complex interactions between and among the facts, data, and findings of our research in terms of costs and benefits to human proficiency and effectiveness in interacting with complex advanced technology and primarily automated systems.

The foregoing comments reflect challenges far beyond the scope of what could be addressed in a few days. However, I propose a matrix structure which may help in focusing our efforts, and perhaps provide an opportunity to make a small contribution to the broader problem.

The triad of training, workload, and performance is so inexorably intertwined and interactive, that to discuss, address, even conduct research on one to the exclusion of the others is at best, irresponsible. Elevated training levels lead to improved performance, and reduced workload may also lead to better performance. What seems to be overlooked is that if the workload is too great during training, not only is performance unacceptable, but learning cannot occur. You cannot learn what you cannot do; after all, the purpose of training is to facilitate learning. But that's only the theoretical prospective. The actual relationship involves the system, which must be operated to a specified performance standard. If the system complexity is high, so will the training requirements be, and the workload experienced by the operator. So, too, can we predict reduced or unreliable performance from the operator. If the system design is made transparent to the operator through an interface which is compatible with the human, the workload will be acceptable and the training requirements manageable, and we can predict satisfactory performance of a reliable nature. The bottom line is that poor or thoughtless system design will lead to poor performance no matter what the training and workload reduction efforts.

Of the many tasks and environments which confront the human operator of complex modern advanced technology systems, there are two which represent extremes in both tasks and environments. Single crew helicopters engaged in night, all weather, nap-of-the-earth military missions and the long-duration, multicrew, space station missions represent these two extremes. These two tasks and their environments, coupled with the triad, make up the matrix I spoke of earlier. The matrix is what I have in mind as a way of focusing our deliberations. The two tasks represent extremes of human interaction with complex modern systems. All else falls in-between.

Barry H. Kantowitz, Department of Psychology, Purdue University, West Lafayette, IN

One important criterion I use for evaluating conferences is the length of my notes. The three presentations for which I had most extensive notes were from Keith Fender, Charles Gainer, and Richard Gabriel. This allows me to infer that for me the most valuable aspect of the conference was the opportunity to interact with colleagues outside of the groves of academe. While the scientific exchanges at the conference were important, I have other ways of obtaining that kind of information--large conferences, a network of informal communication, and journals. I have little opportunity to exchange views with practitioners who have a less theoretical viewpoint due to the urgent demands of immediate problems that must be solved. I was particularly impressed that the practitioners were so aware of the technical problems with their current solutions, and this has increased my motivation to use theoretical tools to help them solve their problems in better ways.

I very much enjoyed the panel interaction and urge that this format be continued. The chance to debate important issues in a small group was quite fruitful. Although my biased view is that the first panel was most diligent, I think all the panels accomplished their goals.

My main complaint about the conference was its length. A full week away from my lab (especially when classes are not in session) is a high price to pay for attendance even at this illuminating conference. I am still trying to catch up. Perhaps a single day of tutorials would have sufficed. Similarly, the panel reports could have been condensed by having fewer panel members make presentations. As is the case with most conferences, a great deal was accomplished over dinner and at night. Thus, the formal scheduled presentations could be condensed without substantially harming the functions of the conference.

Neville Moray, Department of Industrial Engineering, University of Toronto, Ontario, Canada

This was an excellent meeting. It provided an opportunity to clarify my understanding of several topics which are currently of great theoretical importance, such as the status of multiple resource theory and the nature of automatisation of behaviour. The descriptions of the needs of industry, and the opportunity to interact with industry and government representative was unusual and valuable. I found that I have come away with a much clearer picture than before of the extent to which our current knowledge is able to support practical application. I feel that the relation between training, design, and workload was made apparent. In my view training should not be aimed at reducing workload. The combination of design and training should be aimed at providing acceptable performance. Workload measures can be used to guide design and evaluate performance, the aim being to ensure that performance will not require unreasonable effort from the operator. We know enough to put some "envelopes" on design predictively--within the envelope we can expect satisfactory performance. Outside the envelope trouble can be expected. The claim that we can do this was based on the "RFP simulation game" played on the last day.

Penny Sanderson, Department of Psychology, University of Illinois, Urbana-Champaign

The most valuable aspect of the meeting, of course, was the communication which sprang up between the basic-theoretical and applied-practical participants. This was certainly one of the highlights of Panel I's private discussions. The role-playing exercise of Panel IV dramatized the sometimes difficult dialectic of the two sets of participants, and provided us with very good theatre--almost pure psychodrama. As a neophyte in the world of government contracts and so on, I learned a great deal about the process from Panel IV's efforts (and their mutually directed outrage).

In my talk I dealt with the interaction between the validity, reliability, and sensitivity of workload measures. I was suprised at the interest from the applied-practical side in having the test-theoretical constraints further laid out. It seems that a full treatment of this topic would be a useful handbook for practioners.

Another valuable aspect was the discussions and repartee (Neville Moray, Daniel Gopher, Barry Kantowitz, Chris Wickens, etc.) which went on about the heuristic role of theory in guiding research on applied questions. I came away feeling that while theory is inherent in what we do (all observation is theory-laden), there were certainly dangers to external validity in squeezing reality into currently fashionable paradigms.

The Carmel Conference has got me thinking that applied scientists need a good philosophy of applied science. Of course, this is a rather self-refuting cncept, as philosophy is the last thing that practitioners are going to be bothered about. However we do need guidelines about (1) how to manage the trade-off between internal and external validity of experiments and (2) what are the limiting conditions on internal and external validity, and (3) how to identify the subset of experiments or trials which will give the best return with respect to goals agreed upon by researchers and contractors. (I can see the need for this in the NASA Workload, Automation and Training project under way at ARL at the moment, which of course is focused around the LHX just as the conference was.) What would a philosophy of applied science look like? I imagine that along with a large traditional dose of epistemology it would involve aspects of engineering economics, politics, planning, decision-making under risk, pragmatics, and so on. Problem? Maximize the former under constraints provided by the latter. Good topic for a conference?

On the one hand I think the LHX did us a service in providing a good focus for the conference and in being such a borderline piece of equipment from the human factors angle. On the other hand I find it disturbing that such a good question (Interaction of Workload and Training) must be motivated by such a politically conceived piece of equipment. Since the LHX is 90% (pick a number) designed by Pentagon policy dictated by the relative military manpower of the U.S. and U.S.S.R. and the "survivability ratio," I fear our good conference communication is somewhat quixotic.

Jim Staszewski, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA

Let me also thank you and the other organizers for the opportunity to participate in this conference. I came away with many valuable ideas and perspectives, far more than are indicated in the comments below, and I hope that I contributed to the conference in reasonable proportion to what I've gained. I should pass on that every one of the 20 or so participants I spoke with on the meeting's final days expressed similar sentiments as to its value. I interpret this unanimity as an index of unusual success, due unquestionably to the conference's superb organization.

With regard to the publication of the proceedings, I look forward to the opportunity to edit the text of my tutorial presentation and incorporate the appropriate figures. I'm sure it needs more than a little revision if it is to clearly communicate what I'd intended it to.

I've come away from this conference believing that accurate measurement of workload is an unlikely prospect for the immediate future and that definition of this concept may not even be possible. For a concept with such a compelling face validity, this is clearly a frustrating situation. That's the bad news. The good news is that progress has been made in the decade since the 1977 NATO conference, primarily in understanding the complexity of the concept of workload and the corresponding difficulty of defining and measuring it. I think this conference has succeeded in identifying the important theoretical and empirical issues that represent the agenda for basic research on workload and will shape future approaches to its measurement.

With regard to defining workload, as I suggested in my panel presentation, I question whether the concept can be defined. Of course, this view is based on the informal observations of a newcomer to research on this topic, but nevertheless I think a case can be made that workload, like the concept of intelligence, has many of the properties that characterize natural concepts/categories. If this analogy holds, the most relevant of these properties is that the concept has no fixed set of defining features or dimensions that cut across all the situations in which we would want to measure workload.

This conclusion implies that any form of objective, standardized measurement of workload based on a single specified dimension or set of dimensions won't work. Instead, any general approaches to measurement will have to employ instruments that are sensitive to the task-specific nature of workload and be flexible enough to handle a wide range of situations. Perhaps it is because subjective measures have these properties of flexibility and context sensitivity that they enjoy the success Sandy Hart reports. Unfortunately, questions of reliability that pose difficult research problems plague this approach; it's not clear what dimensions people use in making subjective estimates, whether different people use different dimensions (or different weightings of dimensions) to arrive at some composite index, or if people reliably use the same dimensions to estimate the workload of different tasks.

Is there a general approach to measuring workload that is both flexible enough and objective? I think so, and I think that the simulation methodology of cognitive science (the knowledge-engineering approach to task analysis, in Danny Gopher's terms) represents a particularly promising candidate. Generally speaking, this approach is applied to study broad, complex phenomena (such as problem solving, natural language perception and comprehension, memory retrieval, skill acquisition, cognitive development) largely because this methodology can accommodate phenomena in which many endogenous variables (different processes, prior knowledge, capacities, strategies) interact with exogenous variables of the task environment in complex ways that reductionistic, controlled experiments fail to capture. Computer simulation provides a tool for modeling the complex interaction of many variables and also an environment in which these variables can be manipulated experimentally to examine their effects of performance. Applied with sufficient rigor, simulation represents a potentially powerful tool for studying real-world problems, such as workload measurement and the assessment and prediction training effects, problems whose complexity overwhelms traditional experimental analysis.

Some examples of how simulation is being applied to study workload-related issues can be found in recent work on human/computer interaction. The basic strategy is to create a global simulation model with two primary components, a simulation of the particular system under study, and a psychological model of its user. Operator performance can then be examined as the two systems interact in carrying out specified tasks using a specified sequence of procedures. The flexibility of this approach should be apparent. The procedures of the operator could be programmed in various ways to systematically explore the effects of different operator strategies or training-related shifts in operator behavior. Similarly, different device designs can be implemented to examine their effects on performance. From a practical standpoint, the savings achieved by using simulation models of complex systems for testing purposes rather than actual prototypes may well offset costs of this rather expensive and complex approach to workload measurement.

However promising this general approach may be, its actual utility remains an open question and is limited most critically by the accuracy and adequacy which the human user can be modeled. It's clear that viable, psychologically realistic computer models of perception and motor activity are still in the future and, while impressive progress is being made in constructing large-scale, general models of human cognition, these current models rest on far too many arbitrary assumptions and too few empirical facts.

In their defense, however, these systems represent the early and crude prototypes that will guide development of far more sophisticated models. Their psychological validity (or lack thereof) is reflection of the body of sound theory and data available to guide and constrain their design. The point is that cognitive science offers a potentially powerful methodology for measuring workload and analyzing its effects but whether or not this potential is realized depends crucially on advances in our theoretical and empirical understanding of human information processing.

The single issue probably most in need of attention if we are to make progress in understanding, measuring, and predicting workload is that of capacity limitations. That people are limited in their abilities to cope with different tasks is beyond doubt, but widely accepted notions of fixed, fundamental structural or processing limitations can be called into question. Recently, a forceful argument has been made that the entire notion of limited processing resources/capacities is scientifically bankrupt. Empirical studies of digit-span and visual search show that with modest training (by real-world standards) people circumvent what are often assumed to be fundamental structural and processing limitations. While these studies do not refute the existence of structural or processing constraints, they do indicate that the adaptive nature of the human processing system allows it to defeat limiting factors under the appropriate conditions. Without a doubt, a critical item on the research agenda should be identification of the cognitive mechanisms that limit performance and the manner and conditions under which they intervene.

<u>Michael H. Strub, U. S. Army Research Institute, Ft. Bliss, TX</u>

The meeting was an excellent forum for presenting the application needs of the user participants and the state of the art in workload and training research measures/methods.

Presentations dealt with Army MANPRINT needs and ARI aviation workload results comparing one vs. two crewmember's for LHX. Theoretical presentations covered the latest developments in multiple resource attention concepts as well as an assessment of the strenghts and weaknesses of current workload measurment techniques. In the second half of the workshop participants were organized into four task forces to investigate and report back on key aspects of the training-workload interaction.

ARI's increased involvement in operator workload research warrants consideration of joint sponsorship of future workload coordination meetings similar to this one.

<u>Pamela Tsang, NASA Ames Research Center, Moffett Field, CA</u>

Thank you for the opportunity to attend the workshop on Workload and Training. The workshop helped me reconsider some of the "obvious" and often taken for granted issues concerning workload and training.

Much effort had been devoted to defining the concept of workload at the NATO symposium almost a decade ago. The one and only consensus seemed to be that it is a multidimensional concept. While the multidimensionality is perhaps responsible for its unparalleled enthusiasm from diverse sectors, it also made it acceptable for researchers and practioners to define workload convenient only to their theoretical models, research programs, or specific applications. I am therefore questioning the fruitfulness of pursuing standardized workload measures and applying them to training. I would like very much to see some major efforts again to developing a tighter definition of workload. The question of how to measure workload would be secondary to the question of what it is that we are measuring.

<u>Michael Vidulich, NASA Ames Research Center, Moffett Field, CA</u>

To start, I would like to express my appreciation for having been invited to attend the conference. It was, I felt, an exhilarating exchange of ideas.

From my point of view, the most major conclusion of the conference was the consensus that the notion "training reduces workload" is inaccurate. It seems obvious now that no simple relationship exists between training and workload. Given that mental workload is becoming increasingly important as the role of system operators changes from manual control to passive monitoring, it is clear that the relationship between training and workload needs to be better understood. What kind of training reduces workload? Is it possible, or even desirable, for training to reduce workload of all tasks? So far, the most directly applicable research to these questions seems to come from studies of cognitive automaticity. Training that promotes automaticity appears to reduce workload, but mere practice without the development of automaticity may have little effect on workload. However, the development of automaticity may be counterproductive if flexible response to unanticipated situations is required. The workload-related costs and benefits of developing automaticity need to be more fully explored to determine the degree to which they can help define goals and techniques for training programs. Also, theoretical conceptions of mechanisms that could reduce workload without presupposing the development of automaticity would be interesting.

Another major point arrives from the discussion concerning the level of workload that is desirable in a training program. An assumption might be made that training should be performed at a low workload level in order to allow the trainee to learn. While this is certainly true to some degree, it raises an interesting question of when training might profit from higher workload conditions. As one example, the value of training under high stress for what would be a high stress operational situation is an important question.

As this was the first conference to address the issue of the relationship between workload and training, it was inevitable that more questions than answers would be generated. However, I think we can be thankful that so many questions that are both theoretically interesting and practically valuable have been formulated. It seems likely that a great deal of research will be inspired by the conference.

National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-89459		

4. Title and Subtitle	5. Report Date
Workshop on Workload and Training, an Examination of Their Interactions: Executive Summary	July 1987
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Emanuel Donchin (University of Illinois, Urbana-Champaign, IL), Sandra G. Hart, and Earl J. Hartzell	A-87212
	10. Work Unit No.
	505-67-51
9. Performing Organization Name and Address	11. Contract or Grant No.
Ames Research Center Moffett Field, CA 94035	
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Washington, DC 20546	Technical Memorandum
	14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: Sandra G. Hart, Ames Research Center, M/S 239-3, Moffett Field, CA 94035, (415)694-6072 or FTS 464-6072

16. Abstract

This report provides an overview of the Workshop on Workload and Training: An Examination of their Interactions which was held in Carmel, California, from January 5 to 10, 1986. The workshop was jointly sponsored by the Aerospace Human Factors Research Division and the Army Aeroflightdynamics Directorate, and was organized and chaired by Dr. Emanuel Donchin. The goal of the workshop was to bring together experts in the fields of workload and training and representatives from the Department of Defense and industrial organizations who are responsible for specifying, building, and managing advanced, complex systems. The challenging environments and requirements imposed by military helicopter missions and space station operations were presented as the focus for the panel discussions. The workshop permitted a detailed examination of the theoretical foundations of the fields of training and workload, as well as their practical applications. Furthermore, it created a forum where government, industry, and academic experts were able to examine each other's concepts, values, and goals. The discussions pointed out the necessity for a more efficient and effective flow of information among the groups represented. The executive summary describes the rationale of the meeting, summarizes the primary points of discussion, and lists the participants and some of their summary comments. A complete transcription of the tutorials and panel reports is being transcribed and will be published.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Workload Training Helicopters Space	Unclassified - Unlimited Subject Category - 53

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	41	A03

NASA FORM 1626 OCT 86 For sale by the National Technical Information Service, Springfield, Virginia 22161